

Addendum to "Sum rules via large deviations"

Fabrice Gamboa^{*}

Jan Nagel[†]

Alain Rouault[‡]

October 10, 2016

Abstract

In these notes we fill a gap in a proof in Section 4 of Gamboa, Nagel, Rouault [Sum rules via large deviations, J. Funct. Anal. 270 (2016), 509-559]. We prove a general theorem which combines a LDP with a convex rate function and a LDP with a non-convex one. This result will be used to prove LDPs for spectral matrix measures and for spectral measures on the unit circle.

Keywords : Large deviations, random measures

MSC 2010: 60F10

1 Introduction

In Section 4 of [5], we studied large deviations for a pair of random variables with values in topological vector spaces by means of the joint normalized generating function. However, in some cases, the rate function of one of the marginals is not convex, which invalidates this way of proof. Actually, it is possible to state a general theorem which combines a LDP with a convex rate function and a LDP with a non-convex one. It will be used to prove LDPs in [3] for spectral matrix measures and in [4] for spectral measures on the unit circle.

Since this theorem may have its own interest, we give it in a general setting in Section 3

^{*}Université Paul Sabatier, Institut de Mathématiques de Toulouse, 31062 Toulouse Cedex 9, France, e-mail: gamboa@math.univ-toulouse.fr

[†]Technische Universität München, Fakultät für Mathematik, Boltzmannstr. 3, 85748 Garching, Germany, e-mail: jan.nagel@tum.de

[‡]Laboratoire de Mathématiques de Versailles, UVSQ, CNRS, Université Paris-Saclay, 78035-Versailles Cedex France, e-mail: alain.rouault@uvsq.fr

after recalling classical results in Section 2. We come back in Section 4 to the framework of [5].

In the sequel, we assume that \mathcal{X} and \mathcal{Y} are Hausdorff topological vector spaces. \mathcal{X}^* is the topological dual of \mathcal{X} and \mathcal{X} is endowed with the weak topology. We denote by $C_b(\mathcal{Y})$ the set of all bounded continuous functions $\varphi : \mathcal{Y} \rightarrow \mathbb{R}$. A point $x \in \mathcal{X}$ is called an exposed point of a function F on \mathcal{X} , if there exists $x^* \in \mathcal{X}^*$ (called an exposing hyperplane for x) such that

$$(1.1) \quad F(x) - \langle x^*, x \rangle < F(z) - \langle x^*, z \rangle$$

for all $z \neq x$.

2 Some classical results in large deviations

Let us recall two well known results in the theory of large deviations which have to be combined carefully in order to get our general theorem. The first result is the inverse of Varadhan's lemma (Theorem 4.4.2 in [1]), the second one is a version of the so-called Baldi's theorem (Theorem 4.5.20 in [1]). The latter differs from the version in [1] in a straightforward condition to identify the rate function, which was applied for instance in [6] (see also [2]). The proof of our Theorem 3.1 will be quite similar to the proof of these two classical theorems.

Theorem 2.1 (Bryc's Inverse Varadhan Lemma) *Suppose that the sequence (Y_n) of random variables in \mathcal{Y} is exponentially tight and that the limit*

$$\Lambda(\varphi) := \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E} e^{n\varphi(Y_n)}$$

exists for every $\varphi \in C_b(\mathcal{Y})$. Then (Y_n) satisfies the LDP with the good rate function

$$\mathcal{I}(y) = \sup_{\varphi \in C_b(\mathcal{Y})} \{\varphi(y) - \Lambda(\varphi)\}.$$

Furthermore, for every $\varphi \in C_b(\mathcal{Y})$,

$$\Lambda(\varphi) = \sup_{y \in \mathcal{Y}} \{\varphi(y) - \mathcal{I}(y)\}.$$

Theorem 2.2 (A version of Baldi's Theorem) *Suppose that the sequence (X_n) of random variables in \mathcal{X} is exponentially tight and that :*

1. *There is a set $D \subset \mathcal{X}^*$ and a function $G_X : D \rightarrow \mathbb{R}$ such that for all $x^* \in D$*

$$(2.1) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E} \exp(n \langle x^*, X_n \rangle) = G_X(x^*).$$

2. The set \mathcal{F} of exposed points x of

$$G_X^*(x) = \sup_{x^* \in D} \{\langle x^*, x \rangle - G_X(x^*)\},$$

with an exposing hyperplane x^* satisfying $x^* \in D$ and $\gamma x^* \in D$ for some $\gamma > 1$, is dense in $\{G_X^* < \infty\}$.

Then (X_n) satisfies the LDP with good rate function G_X^* .

3 A general theorem

Our extension is the following combination of the two above theorems. The main point is that the rate function does not need to be convex, but we still only need to control linear functionals of X_n .

Theorem 3.1 Assume that $X_n \in \mathcal{X}$ and $Y_n \in \mathcal{Y}$ are defined on the same probabilistic space. Moreover, we assume that the two sequences (X_n) and (Y_n) are exponentially tight. Assume further that :

1. There is a set $D \subset \mathcal{X}^*$ and functions $G_X : D \rightarrow \mathbb{R}$, $J : C_b(\mathcal{Y}) \rightarrow \mathbb{R}$ such that for all $x^* \in D$ and $\varphi \in C_b(\mathcal{Y})$

$$(3.1) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E} \exp (n \langle x^*, X_n \rangle + n \varphi(Y_n)) = G_X(x^*) + J(\varphi).$$

2. The set \mathcal{F} of exposed points x of

$$G_X^*(x) = \sup_{x^* \in D} \{\langle x^*, x \rangle - G_X(x^*)\},$$

with an exposing hyperplane x^* satisfying $x^* \in D$ and $\gamma x^* \in D$ for some $\gamma > 1$, is dense in $\{G_X^* < \infty\}$.

Then, the pair (X_n, Y_n) satisfies the LDP with speed n and good rate function

$$\mathcal{I}(x, y) = G_X^*(x) + \mathcal{I}_Y(y),$$

where

$$\mathcal{I}_Y(y) = \sup_{\varphi \in C_b(\mathcal{Y})} \{\varphi(y) - J(\varphi)\}.$$

Let us note that in view of Varadhan's Lemma we have

$$J(\varphi) = \sup_{y \in \mathcal{Y}} \{\varphi(y) - \mathcal{I}_Y(y)\}.$$

Proof:

Upperbound: The proof follows the lines of the proof of part (b) of Theorem 4.5.3 in [1]. Note that since the sequence (X_n, Y_n) is exponentially tight it suffices to show the upper bound for compact sets.

Lowerbound: As usual, it is enough to consider a neighbourhood $\Delta_1 \times \Delta_2$ of (x, y) where $\mathcal{I}(x, y) < \infty$. Take $\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}((X_n, Y_n) \in \Delta_1 \times \Delta_2)$ and get a lower bound tending to $\mathcal{I}(x, y)$ when the size of the neighbourhood tends to zero. Actually, due to the density assumption 2. it is enough to study the lower bound of $\mathbb{P}(X_n \in \Delta_1, Y_n \in \Delta_2)$ when $x \in \mathcal{F}$ and $\mathcal{I}_Y(y) < \infty$.

As in [1] (Proof of Lemma 4.4.6), let $\varphi : \mathcal{Y} \rightarrow [0, 1]$ be a continuous function, such that $\varphi(y) = 1$ and φ vanishes on the complement Δ_2^c of Δ_2 . For $m > 0$, define $\varphi_m := m(\varphi - 1)$. Note that

$$J(\varphi_m) \geq -\mathcal{I}_Y(y).$$

We have

$$\mathbb{P}(X_n \in \Delta_1, Y_n \in \Delta_2) = \mathbb{E} \left[\mathbb{1}_{\{X_n \in \Delta_1\}} \mathbb{1}_{\{Y_n \in \Delta_2\}} e^{n\langle x^*, X_n \rangle + n\varphi_m(Y_n)} e^{-n\langle x^*, X_n \rangle - n\varphi_m(Y_n)} \right].$$

Now $-\varphi_m \geq 0$ and on Δ_1 , $-\langle x^*, X_n \rangle \geq -\langle x^*, x \rangle - \delta$ for a $\delta > 0$, so that

$$(3.2) \quad \mathbb{P}(X_n \in \Delta_1, Y_n \in \Delta_2) \geq \mathbb{E} \left[\mathbb{1}_{\{X_n \in \Delta_1\}} \mathbb{1}_{\{Y_n \in \Delta_2\}} e^{n\langle x^*, X_n \rangle + n\varphi_m(Y_n)} \right] e^{-n\langle x^*, x \rangle - n\delta}.$$

Denoting

$$\ell_n = \frac{1}{n} \log \mathbb{E} e^{n\langle x^*, X_n \rangle}, \quad \mathcal{L}_n := \frac{1}{n} \log \mathbb{E} e^{n\langle x^*, X_n \rangle + n\varphi_m(Y_n)}$$

and $\tilde{\mathbb{P}}$ the new probability on $\mathcal{X} \times \mathcal{Y}$ such that

$$\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} = e^{n\langle x^*, X_n \rangle + n\varphi_m(Y_n) - n\mathcal{L}_n},$$

we get

$$(3.3) \quad \mathbb{P}(X_n \in \Delta_1, Y_n \in \Delta_2) \geq \tilde{\mathbb{P}}(X_n \in \Delta_1, Y_n \in \Delta_2) e^{-n\langle x^*, x \rangle - n\delta + n\mathcal{L}_n}.$$

For the exponential term we have

$$(3.4) \quad \liminf_{n \rightarrow \infty} \frac{1}{n} \log e^{-n\langle x^*, x \rangle - n\delta + n\mathcal{L}_n} \geq \langle x^*, x \rangle - \delta + G_X(x^*) + J(\varphi_m) \geq -G_X^*(x) - \mathcal{I}_Y(y) - \delta.$$

We may choose δ arbitrarily small by choosing Δ_1 sufficiently small, so that it will be enough to prove that

$$(3.5) \quad \tilde{\mathbb{P}}(X_n \in \Delta_1, Y_n \in \Delta_2) \xrightarrow[n \rightarrow \infty]{} 1$$

or equivalently, that

$$(3.6) \quad \tilde{\mathbb{P}}(X_n \in \Delta_1^c) + \tilde{\mathbb{P}}(Y_n \in \Delta_2^c) \xrightarrow{n \rightarrow \infty} 0.$$

For the first term, note that under $\tilde{\mathbb{P}}$ the moment generating function of X_n satisfies

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \log \tilde{\mathbb{E}}[e^{n\langle z^*, X_n \rangle}] &= \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E}[e^{n\langle z^* + x^*, X_n \rangle + \varphi_m(Y_n) - n\mathcal{L}_n}] \\ &= G_X(z^* + x^*) + J(\varphi_m) - G_X(x^*) - J(\varphi_m) \\ &= G_X(z^* + x^*) - G_X(x^*) \\ &=: \tilde{G}_X(z^*), \end{aligned}$$

for $z^* \in \tilde{D} := \{z^* : x^* + z^* \in D\}$. We may then follow the argument on p.159-160 in [1] (as an auxiliary result in their proof of the lower bound). Using that $x^* \in D$ is an exposing hyperplane, we get

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \tilde{\mathbb{P}}(X_n \in \Delta_1^c) < 0.$$

Considering the second term in (3.6), we have, on Δ_2^c

$$\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} = e^{-nm + n\langle x^*, X_n \rangle - n\mathcal{L}_n}$$

so that

$$\tilde{\mathbb{P}}(Y_n \in \Delta_2^c) \leq e^{-nm + n\ell_n - n\mathcal{L}_n}.$$

Taking the logarithm, this implies

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{n} \log \tilde{\mathbb{P}}(Y_n \in \Delta_2^c) &\leq -m + G_X(x^*) - G_X(x^*) - J(\varphi_m) \\ &= -m - \sup_{z \in \mathcal{Y}} \{\varphi_m(z) - \mathcal{I}_Y(z)\} \leq -m + \mathcal{I}_Y(y) \end{aligned}$$

which tends to $-\infty$ when $m \rightarrow \infty$.

To summarize, we have proved (3.6), i.e. (3.5), which with (3.3) and (3.4) gives

$$\lim_{\Delta_1 \downarrow x, \Delta_2 \downarrow y} \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(X_n \in \Delta_1, Y_n \in \Delta_2) \geq -G_X^*(x) - \mathcal{I}_Y(y),$$

which leads to the lower bound of the LDP. \square

4 Joint LDP for measure and truncated eigenvalues

In Section 4 of [5], we studied the joint moment generating function of a non-negative measure $\tilde{\mu}_{I(j)}^{(n)}$ on a compact set $[\alpha^-, \alpha^+]$ and a collection of j extremal support points

$\lambda_M^\pm(j) \in \mathbb{R}^{2j}$, restricted to the compact set $[-M, M]$. For the sake of a clearer notation, we drop here the dependency on j . It is shown in Theorem 4.1 in [5], that λ_M^\pm satisfies the LDP with speed n and good rate $\mathcal{I}_{M,\lambda^\pm}$. Furthermore, the sequence of $\tilde{\mu}_I^{(n)}$ is exponentially tight and if

$$\mathcal{G}_n(f, s) = \mathbb{E} \left[\exp \left\{ n \int f d\tilde{\mu}_I^{(n)} + n \langle s, \lambda_M^\pm \rangle \right\} \right],$$

then for all f such that $\log(1 - f)$ is continuous and bounded and all $s \in \mathbb{R}^{2j}$,

$$(4.1) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathcal{G}_n(f, s) = G(f) + H(s),$$

with G^* strictly convex on a set of points dense in $\{G^* < \infty\}$.

However, the rate $\mathcal{I}_{M,\lambda^\pm}$ might be non-convex and hence the dual H^* is not strictly convex on a dense set. The convergence in (4.1) is therefore not enough to conclude the joint LDP for $(\tilde{\mu}_I^{(n)}, \lambda_M^\pm)$ directly from the classical Theorem 2.2.

To show the joint LDP, we will apply Theorem 3.1. Indeed, let D be the set of bounded continuous functions f from $[\alpha^-, \alpha^+]$ to \mathbb{R} such that $\sup_x f(x) < 1$. If we define for $\varphi : \mathbb{R}^{2j} \rightarrow \mathbb{R}$ continuous and bounded and $f \in D$

$$\hat{\mathcal{G}}_n(f, \varphi) = \mathbb{E} \left[\exp \left\{ n \int f d\tilde{\mu}_I^{(n)} + n \varphi(\lambda_M^\pm) \right\} \right],$$

then the same arguments as in Section 4 of [5] show

$$(4.2) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \log \hat{\mathcal{G}}_n(f, \varphi) = G(f) + J(\varphi),$$

where

$$G(f) = - \int \log(1 - f) d\mu_V$$

for a probability measure μ_V . Moreover, in Section 4 of [5] it is shown that every measure on $[\alpha^-, \alpha^+]$ with a strictly positive continuous density h with respect to μ_V is an exposed point and the exposing hyperplane is the function $f = 1 - h^{-1}$. Since f is continuous and strictly less than 1, $\gamma f \in D$ for $\gamma > 1$ small enough. The assumptions of Theorem 3.1 are then satisfied, which yields the LDP for $(\tilde{\mu}_I^{(n)}, \lambda_M^\pm)$ with good rate

$$\mathcal{I}(\mu, \lambda) = G^*(\mu) + \mathcal{I}_{M,\lambda^\pm}.$$

After taking the limit $M \rightarrow \infty$, this proves the statement of Theorem 4.2 of [5].

References

- [1] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*. Springer, 1998.
- [2] H. Dette and F. Gamboa. Asymptotic properties of the algebraic moment range process. *Acta Mathematica Hungarica*, 116:247–264, 2007.
- [3] F. Gamboa, J. Nagel, and A. Rouault. Sum rules and large deviations for spectral matrix measures. arXiv preprint arXiv:1601.08135, 2016.
- [4] F. Gamboa, J. Nagel, and A. Rouault. Sum rules and large deviations for spectral measures on the unit circle. arXiv preprint arXiv: 1604.06934, 2016.
- [5] F. Gamboa, J. Nagel, and A. Rouault. Sum rules via large deviations. *J. Funct. Anal.*, 270(2):509 – 559, 2016.
- [6] F. Gamboa, A. Rouault, and M. Zani. A functional large deviation principle for quadratic forms of gaussian stationary processes. *Stat. and Probab. Letters*, 43:299–308, 1999.